

NMED Data Summaries

Nutrient Data Reduction

The NMED nutrient dataset was compiled from multiple files provided by NMED (Table 1). These files were compiled in a relational database, along with information regarding site characteristics, non-nutrient water quality, chlorophyll a, macroinvertebrate samples, and diatom samples. Similar data were compiled for other data sources, including the National Rivers and Streams Assessment (NRSA) and the Wadeable Streams Assessment (WSA).

Table 1. NMED data sources for nutrient analysis in NM streams.

File Name	Date Range	No. Records	Remarks
Master_Streams_N+N_4_18.xls	1990-2006	3743 samples, 708 sites	NO3NO2 only, 179 records from before 1998
Master_Streams_TKN_4_18.xls	1990-2006	4187 samples, 757 sites	TKN only, 583 records from before 1998
Master_Streams_Phosph_4_18	1990-2006	7215 samples, 979 sites	TP only, 3324 records from before 1998
CHEMLABNutsdata.xlsx	1999-2009	5142 samples, 802 sites	Data for major nutrients mostly complete. Several records duplicated in "Master_" files.
RiverNutrients_20120830.xlsx	2009 - 2012	948 samples, 255 sites	Data for major nutrients mostly complete.
Benthic Nutrient Data.xlsx	1994-1999	15 samples, 6 sites	Nutrient and other common analytes mostly complete
REMAP Chem.xlsx	1999-2001	49 samples, 40 sites	Nutrient and other common analytes mostly complete

Detection limits were identified for NO3NO2, TKN, and TP that was not necessarily as reported, but represented a substantial "shelf" in the data sets. For TP it was 0.03 mg/L and for NO3NO2 and TKN, it was 0.1 mg/L. All values at or below these standard values were re-set to half the value (0.015 and 0.05 mg/L). A few (1-10) high outliers were removed for each nutrient species. They were assumed to be recorded in error because they were much higher than other values in the entire dataset and, in most cases, in comparison to other samples at the same sites. No systematic unit errors were suspected, except for five NO3NO2 values, which appeared to be µg/L instead of mg/L. These values were removed instead of converted by assumed units. Total N (TN_calc) was calculated as the sum of TKN + NO3NO2. In general, high TP values were correlated with high TN_calc values and TN_calc values were typically higher than TP values (Figure 1). Nutrient values were averaged within each year-month. In the first figure, values that fall between half the standard detection value and the detection value result from averaging a higher value with a ND value.

Comment [FJ1]: This may prove troublesome depending on the actual number of non-detects.

Comment [FJ2]: I am not sure whether I understand this statement but will need to follow up with Ben.

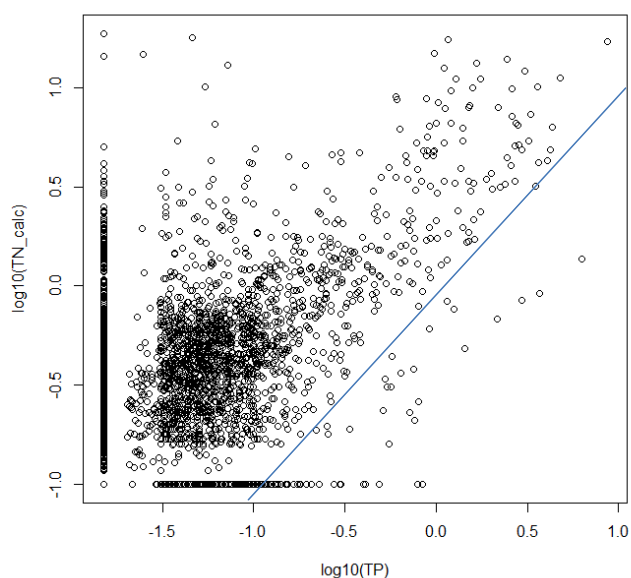


Figure 1. Distribution of log transformed TN and TP values in sites with samples in multiple months/year, showing the 1:1 line.

For an analyses of temporal variability, only site-years with more than two samples were used. This was done so that variation among months could be attributed to changes within sites, reducing the effects among sites. There were more TP values than TN values in the reduced dataset (Table 2). Variability in the data was calculated from an ANOVA grouping data by site-year and calculating the root mean squared error (RMSE) from the ANOVA MSE. This is analogous to an average standard deviation within groups. For TN, the RMSE is 0.95 mg/L, or 1.6 times the mean of 0.59 for this data set. For TP, the RMSE is 0.29 mg/L, or 2.6 times the mean of 0.11 for this data set. Most samples were collected between March and October in the years 1990-2012 (Table 3).

Table 2. Summary statistics for the reduced data set used in temporal nutrient analysis.

	N	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	RMSE
TP	4849	0.015	0.015	0.032	0.1092	0.07	8.65	0.29
TN_calc	3733	0.1	0.17	0.302	0.5929	0.54	18.74	0.95

Table 3. Distribution of TP and TN samples by month of the year.

	1	2	3	4	5	6	7	8	9	10	11	12
TP	71	73	374	573	662	451	691	537	482	654	202	79
TN_calc	15	34	263	442	544	385	524	447	369	537	126	47

When plotting all data in yearly categories, it is apparent that there is a shift in values starting in 1998 (Figure 2). For TN, all values before 1998 were < 0.25 mg/L in this reduced data set and < 3.5 mg/L in the complete data set. After 1998, the values were higher, up to 18.7 mg/L (mean 0.65 mg/L). For TP, it appears that there was a change in the detection limits after 1998. Before 1998, 62% of measures were below the standardized ND level, compared to 39% after 1998. The mean of values before 1998 was somewhat lower than after (0.095 mg/L compared to 1.1 mg/L).

Comment [FJ3]: See comment FJ1

If we accept that the data are somehow biased before 1998, we can examine seasonal nutrient values using box plots by month. When samples were limited to those collected after 1997, it appeared that samples collected in January have higher values of TN and TP. In addition, samples collected in December had higher TN and more frequent non-detects in TP. Frequent non-detects (> 50% of records) in TP also occurred in February and November.

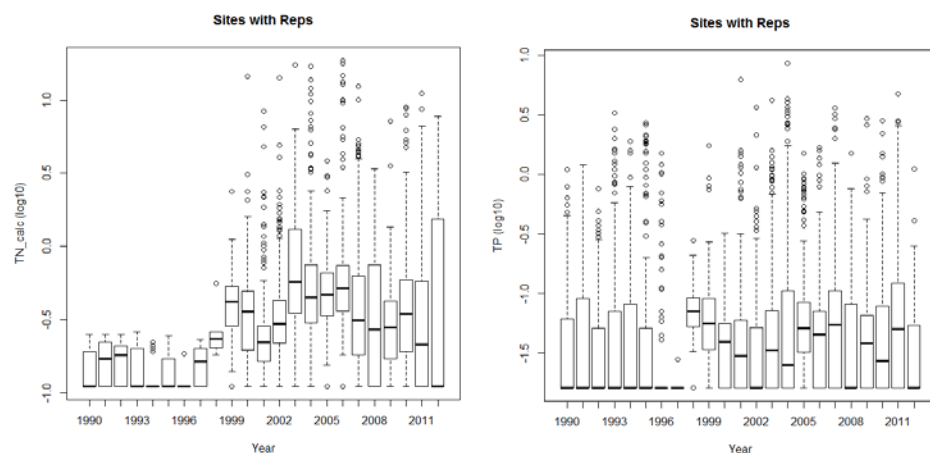


Figure 2. Distributions of TN and TP values by year, showing that before 1998, TN values were consistently lower and there were more numerous TP non-detect values.

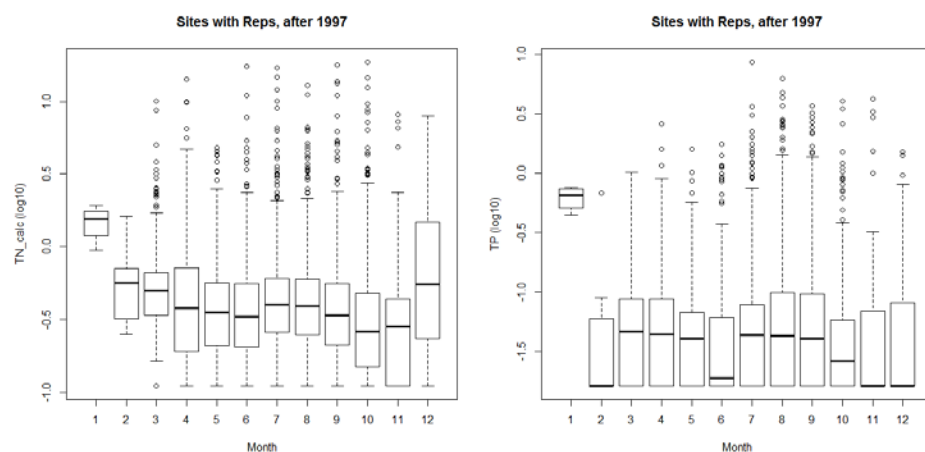


Figure 3. Distributions of TN and TP values collected after 1997 by month, showing that TN values are generally higher in January and December, TP values are higher in January, and more non-detect TP values are in February, November, and December.

Chlorophyll *a*

Chlorophyll *a* data for benthic samples were provided by NMED in two files (Table 4). One sample, collected in March, was excluded because all other samples were collected August – November. After removing sites designated “_CMP”, 235 chlorophyll *a* samples from 181 sites were associated with nutrient data collected within 30 days. Distributions of samples among months and among years did not reveal any biased values or trends. Some sites (~58) were designated as “rivers”, and they may be removed from some analyses, but were included in the following statistics.

Comment [FJ4]: We should be excluding rivers since this is beyond the scope of the N-Steps proposal.

Table 4. Chlorophyll *a* data

File Name	Date Range	No. Records	Remarks
NM_ChlorophyllData_20120918.xlsx	2004 - 2011	273 samples, 208 sites	Some samples (~58) are from large rivers, which may be excluded from some analyses
RiverNutrients_20120830.xlsx	2011	25 samples	Redundant with data in NM_ChlorophyllData_file

For a correlation analysis on untransformed data, one randomly selected sample per site was used. Significant positive Pearson correlations were found with TN, NO₃NO₂, and TP ($p < 0.05$). However, when Spearman ranked correlations were calculated, only conductivity and DO were significant (Table 5). This suggests that untransformed high outlier values and several low values for chlorophyll *a* and nutrients (Table 6) have an effect on the Pearson correlations. After log-log transformation, the relationship

between nutrients and chlorophyll *a* is not strong (Figure 4). Based on 40 sites with multiple samples (57 samples), the RMSE for chlorophyll *a* within sites over multiple years was calculated as 4.65 $\mu\text{g}/\text{cm}^2$.

Comment [FJ5]: Ecoregion? and is this disassociation due to sediments?

Table 5. Correlation coefficients with chlorophyll *a* concentrations. Significant coefficients are marked with an asterisk (*: $p < 0.05$; **: $p < 0.01$).

	Pearson	Spearman
TN_calc	0.20**	0.04
NO3NO2	0.20**	0.03
TKN	0.07	0.01
TP	0.25**	0.07
Ammonia	0.09	-0.04
pH	-0.10	0.04
Conductivity	0.00	0.26**
Temperature	-0.07	-0.06
DO	0.14	0.16*
DOsat	0.05	0.05
Turbidity	-0.10	0.07

Table 6. Summary statistics for the reduced data set used for chlorophyll *a* analysis.

	N	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Chl_a	191	0.02	1.25	3.48	5.96	7.04	55.47
TP	176	0.015	0.015	0.034	0.109	0.085	2.2
TN_calc	176	0.1	0.238	0.38	0.793	0.685	18.26

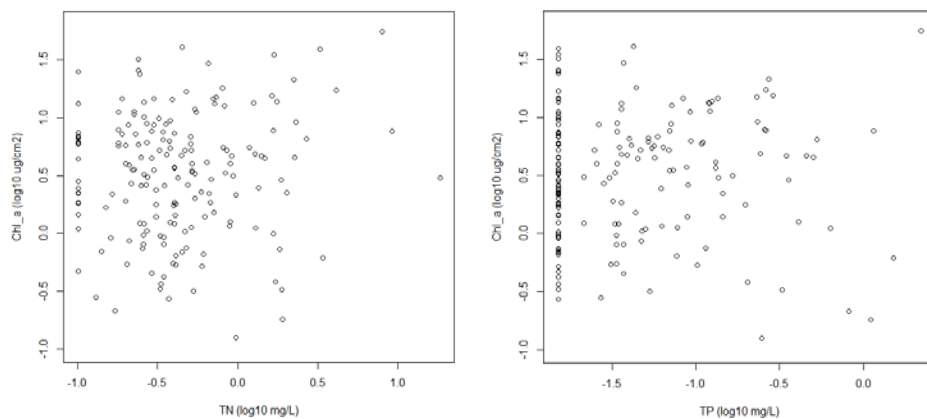


Figure 4. Chlorophyll *a* in relation to TN and TP (log-log transformations).

In the NRSA data, 88 sites with chlorophyll *a* data from periphyton or water, or both, were available. Half of the sites are in NM and the other half are from surrounding plains, mountains, or xeric areas.

Dissolved Oxygen

NMED provided dissolved oxygen (DO) statistics for diel data records over 206 time periods in 188 sites. The number of valid samples associated with nutrient data (within 30 days of the DO start date) was reduced to 148 samples in 141 sites. The time periods included several day/night cycles with frequent (at least hourly) DO readings. Rates of productivity and respiration were estimated from the differences in DO concentrations over 2, 3, and 4 hours (Table 7). The maximum positive difference in DO for each 4 hour time period over several days was an estimate of maximum productivity. The maximum negative difference was an estimate of maximum respiration. The 4 hour time period is recommended because it is long enough to average out possible anomalies that may occur in the shorter periods and it is similar to time periods used by the USGS (Colombo et al. 2004).

Table 7. Productivity (P) and respiration (R) statistics for all samples

	2hrPmax	3hrPmax	4hrPmax	2hrRmax	3hrRmax	4hrRmax
Min.	0.02	0.01	0.01	-6.82	-5.5	-4.71
1 st . Quartile	0.19	0.16	0.16	-0.78	-0.67	-0.60
Median	0.40	0.33	0.28	-0.46	-0.39	-0.36
Mean	0.60	0.49	0.43	-0.66	-0.55	-0.49
3 rd . Quartile	0.75	0.64	0.53	-0.30	-0.26	-0.24
Max.	6.5	4.54	3.65	-0.07	-0.05	-0.05

Spearman correlations between the production and respiration statistics and nutrient variables were not strong (Table 8). Correlations with water quality variables were stronger, suggesting that the water quality environment has a stronger effect on the DO patterns than the nutrients. Some of these effects may be accounted for with site classification or partial correlations.

Table 8. Spearman correlation coefficients between productivity (P) and respiration (R) statistics and nutrient and water quality variables.

Nutrient Variable	4hrPmax	4hrRmax	Water Quality Variable	4hrPmax	4hrRmax
TP	0.14	-0.16*	Chlorophyll a	0.19*	-0.15
TN_calc	0.19*	-0.16*	pH	0.22**	-0.26**
NO3NO2	0.10	-0.05	conductivity	0.34***	-0.26**
TKN	0.07	-0.07	temperature	0.29***	-0.24**
Ammonia	-0.07	0.07	DO grab sample	0.22**	-0.25**

Benthic Macroinvertebrates

837 benthic samples were selected for analysis, and were associated with sites named using NMED conventions, 1990-2011. 323 samples could not be associated with nutrient chemistry in the same year as the benthic sample. These sites were originally identified for inclusion in the analyses because both a benthic sample and a nutrient sample were associated with the site (regardless of timing).

In a concurrent NMED data set, 386 samples (in 251 sites) were associated with nutrient data collected within 30 days of the benthic sampling date. In an annually summarized data set, 491 samples (in 289 sites) were associated with nutrient data collected in March – October of the same calendar year as the benthic sampling date (Table 9). 114 sites of the 491 have multiple samples either over time or by sampling method. Samples collected by NRSA (N=44) and WSA (N = 87) can be included in analyses. For the WSA data, the original taxa lists are available and metrics can be calculated alongside NMED data. Fewer than 87 sites are valid for analysis due to replicates within sites and a few samples outside of the study area. For the NRSA data, three methods were used, depending on site gradient and waterbody size. We only have metrics for NRSA data.

Benthic metrics were checked for consistency among sites in the annually summarized data set (491 samples). For pooling data in analyses, total numbers of individuals and total numbers of taxa should not show bias among methods or over years. When numbers of individuals and taxa counts were compared among collection methods, there were few methods that had obvious bias (Figure 5). Of the more common methods, the distributions of values were similar, except that the Reachwide method appeared to produce more taxa than the Jacobi Hess method. Some of the less common methods produced lower counts of taxa (Low Gradient Wadeable [BELGW], Low Gradient Boatable [BELGB], Large River, and Surber) and lower numbers of individuals (Low Gradient Wadeable).

The number of individuals has an effect on the taxa count, especially for small samples (Figure 6). When samples have less than 100 organisms, the taxa count obviously lower than the taxa counts for samples with 300 organisms or more. Very large samples do not have increasingly higher taxa counts.

Comment [FJ6]: Need to follow-up with Ben since for low gradient streams, two benthic samples were collected. I don't understand what the last sentence means.

Table 9. Numbers of benthic samples by sampling program, collection method, and years.

Yr.	NMED								NRSA			EMAP WSA	
	<u>Canton Hess</u>	<u>Jacobi Hess</u>	<u>Surber</u>	<u>KickNet</u>	<u>Unknown</u>	<u>EMAP Large River</u>	<u>EMAP Reachwide</u>	<u>EMAP-TR</u>	<u>Reachwide</u>	<u>Low Gradient</u>	<u>Boatable</u>	<u>Reachwide</u>	<u>Targeted Riffle</u>
'90		28		3									
'91		7		3									
'92		23	6	5									
'93		18		1									
'94		5		2									
'95		1											
'96			2										
'97		4											
'98		3		19									
'99		4			3	1	19						
'00		9		20			21					10	9
'01	3	12		12			3					11	7
'02	9			7								12	5
'03	1	4			2							10	10
'04	29		5		6		1					12	
'05	28	1	3	5									
'06				2			16	12					
'07							31	28					
'08				3					16	2			
'09				20					13		13		
'10				17									
'11				17									

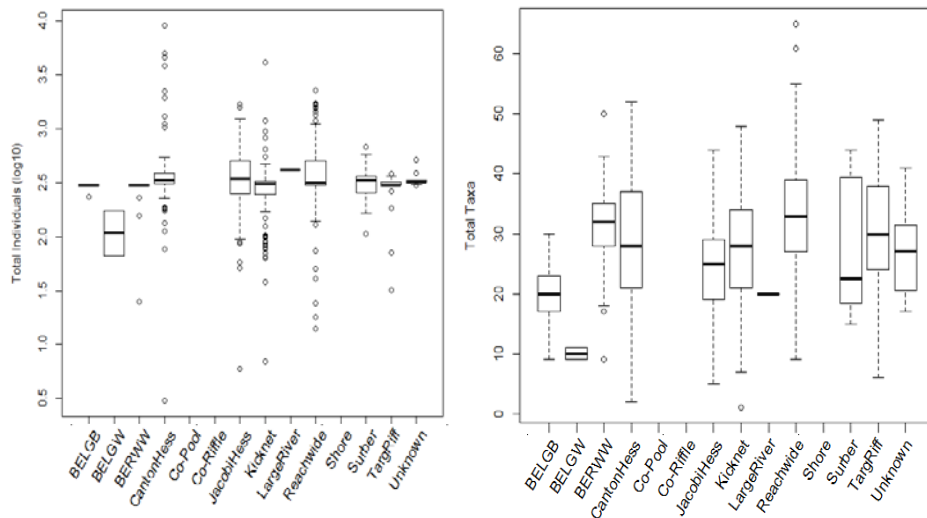


Figure 5. Numbers of individuals and taxa counts by collection method.

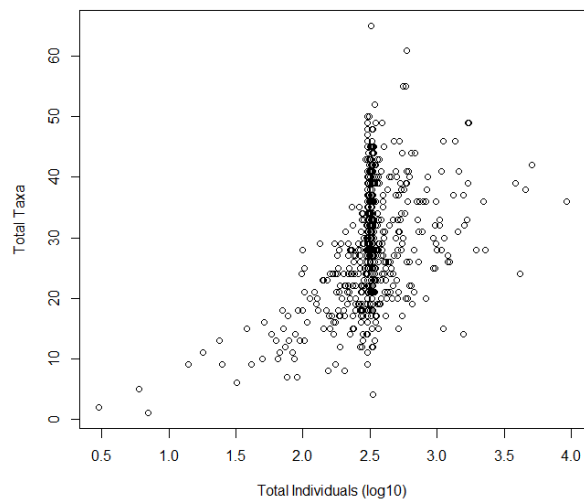


Figure 6. Numbers of individuals compared to taxa counts in all methods combined.

When the numbers of individuals and taxa counts were compared over time, it was apparent that samples collected before 1999 were somewhat different than those collected after 2000 (Figure 7). The early samples had more variable numbers of individuals, but with narrower overall ranges. They also had lower taxa counts in general than the later samples. There were only two collection methods represented by substantial numbers of sample before and after 1999; the Jacobi Hess and the Kicknet samples. All other sample methods were only used after 1999, except for a few Surber samples. When comparing over time within methods, it was apparent that the Jacobi Hess samples had similar distributions over time. However, early Kicknet samples were more variable in size with fewer taxa than later samples.

In an investigation of taxonomic resolution over time for kicknet samples, there were taxa groups that became more diverse in 1998 and after, suggesting a different taxonomic target or refined taxonomist expertise. In the later time period, 315 new taxa were identified. Specific groups that were more diverse in later years included Oligochaeta, Arachnida, Ceratopogonidae, and the Chironomidae (especially the genera Eukiefferiella, Orthocladus, and Cricotopus. The Jacobi-modified Hess samples collected before 1998 were similar to the early kicknet samples in terms of taxa identified. There was not an obvious shift in taxonomic resolution over time in the Jacobi-modified Hess samples.

Taxonomic resolution among methods was relatively consistent in the samples collected after 1997. There were specific taxa groups where the resolution differed among methods, such as Oligochaeta, Trombidiformes (Arachnida), and the chironomid genera Eukiefferiella and Orthocladus, which were identified only to genus in the reachwide and targeted riffle samples. These differences were not substantial enough to warrant taxonomic re-identification.

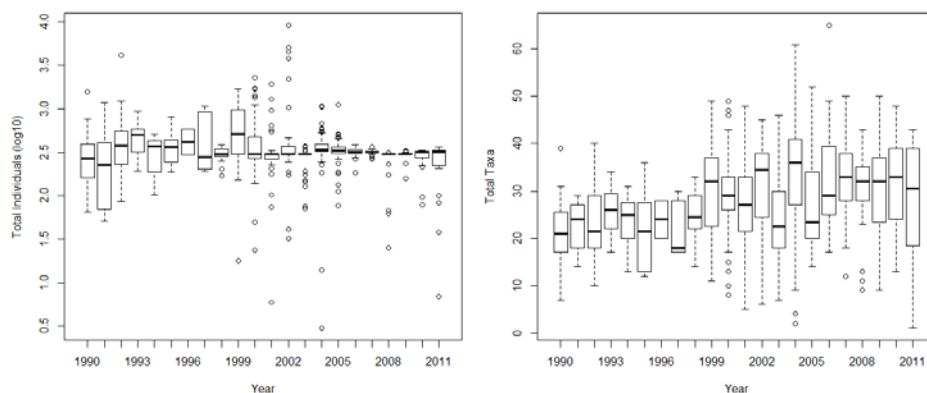


Figure 7. Numbers of individuals and taxa counts by year of collection.

Diatoms

In a NMED data set, 295 samples (in 203 sites) were associated with nutrient data collected within 30 days of the diatom sampling date (in years 2002 – 2010). Additional samples were not readily associated with nutrient data. Samples typically included 500-600 valves and between 20 and 70 taxa. 78 metrics were calculated using taxonomic hierarchies and attributes from multiple sources that are relevant in the western US. The attributes used in calculating these metrics were not complete for all taxa. Only 31 metrics had attributes associated with more than 40% of the taxa.

In the NRSA data, calculated metrics for 68 samples from all site types (wadeable, boatable, low gradient) were available for analysis. Raw data are also available. Samples typically included 600 valves and between 33 and 95 taxa. Because of differences in sample processing and taxonomic identifications, pooling NMED and NRSA diatom data in analyses should be further investigated. The taxonomic attributes for the NRSA taxa have been requested. Reconciliation of taxa and attributes among data sets will be required for comparable metric calculation.

Comment [FJ7]: Follow up with Ben regarding the specifics and the issue with low gradient samples.

Assemblage metrics are available for both NMED and NRSA data sets, though they are not as yet based on identical taxa attributes or calculation conventions. Reconciliation of attributes and calculation of the NRSA metrics as they are already done for NMED is planned. The NRSA developed an index for diatoms that would be applicable in NM. Application of the NRSA index may be possible after reconciling the NMED taxa to the NRSA attributes.

Conclusions

Subsets of the data should be considered for some analyses. Our investigation suggests that samples collected before 1998 had different distributions of values, especially for TN. Possible causes for the differences may include analysis using different laboratory methods or targeting of different site types. NMED confirmed that there was a change in the detection limits for samples processed before and after 1998. They also confirmed that sites were not targeted differently and sampling methods did not change during those time periods. Mixing samples from before 1998 with those after may give uninterpretable results for certain analyses. For TP, the differences are related more to low values than to high values and analyses that exclude non-detects may be applicable in all years.

Seasonality of nutrient concentrations may also be in effect. The non-winter months appear to have more consistent signals and they may be better suited to stressor-response analyses. Considering that samples are most numerous in the months of March – October, there may be justification for limiting some analyses to the common and consistent sampling period. We also looked at seasonality in a subset of sites – only those that had winter samples. By eliminating samples that were only sampled in non-

winter months, we removed the possibility that the sampling design was biased against winter samples. In these analyses, the patterns were similar to those observed in the larger data set.

With RMSE values greater than the annual mean, there is some justification for incorporating more than individual grab sample nutrient results for comparisons to metrics of chlorophyll *a*, macroinvertebrates, diatoms, or dissolved oxygen. Because the winter samples are less common and somewhat biased in value, the nutrient statistics for comparison to the response variables should include the nutrient sample within one month of the response measure (the grab sample), the mean of monthly averages from March through October, and the maximum value during those months.

The weak correlations between nutrients and chlorophyll *a* and between nutrients and dissolved oxygen may be due to as yet unanalyzed factors, such as canopy cover, substrate, flow, turbidity, or other elusive measures. Classification into natural stream types and partial correlation techniques should be further investigated.

Benthic samples with fewer than 100 organisms should be excluded from analyses. There is a possibility of pooling benthic samples for analysis from multiple sampling programs and sampling protocols. Benthic samples collected before 1999 and with certain methods might be excluded or isolated to reduce variability introduced by sampling protocols. It is recommended that samples collected before 1998 from all methods can be pooled for analysis, but kept separate from later samples. After 1997, samples collected using Jacobi-modified Hess samples and low gradient wadeable and boatable methods should be removed or further scrutinized. The NRSA program data were not included in the benthic database and metrics were therefore calculated using different algorithms and probably different taxonomic attributes than those in the NMED database. Complex and attribute-based metrics from NRSA may not be directly comparable to the NMED and WSA metrics. Further investigation for pooling samples would require ordination or cluster analysis to determine the compositional similarity of samples, which would be more informative than number of individuals and taxa counts alone.

References Cited

Colombo, M.J. S.J. Grady, and E.C. Todd Trench. 2004. Nutrient Enrichment, Phytoplankton Algal Growth, and Estimated Rates of Instream Metabolic Processes in the Quinebaug River Basin, Connecticut, 2000–2001. Scientific Investigations Report 2004-5227. U.S. Geological Survey, in cooperation with the Connecticut Department of Environmental Protection.
<http://pubs.usgs.gov/sir/2004/5227/sir2004-5227.pdf>